Learning a Predictive Model of Human Gait for the Control of a Lower-limb Exoskeleton

Erwin Aertbeliën and Joris De Schutter

Abstract— **For an intelligent dynamic motion interaction between a human and a lower-limb exoskeleton, it is necessary to predict the future evolution of the joint gait trajectories and to detect which phase of the gait pattern is currently active. A model of the gait trajectories and of the variations on these trajectories is learned from an example data set. A gait prediction module, based on a statistical latent variable model, is able to predict, in real-time, the future evolution of a joint trajectory, an estimate of the uncertainty on this prediction, the timing along the trajectory and the consistency of the measurements with the learned model. The proposed method is validated using a data set of 54 trials of children walking at three different velocities.**

I. INTRODUCTION

An intelligent dynamic motion interaction between a human and a robotic assistive device, such as a lower-limb exoskeleton, calls for advanced interaction modes such as assistance-as-needed in view of physiotherapeutic objectives (e.g. training, rehabilitation), adaptation to the human (e.g. variation in time), recognition of the human intent (e.g. to perform different tasks), etc. A key component in such a system is a gait prediction module, i.e. a module that can estimate continuously the movements of the subject wearing the exoskeleton and that can anticipate future movements.

In contrast to methods that start from a model, such as [1], or methods that start from a force field and/or a "virtual tunnel" [2], this paper learns a model of the gait trajectories from a data set of example gait patterns during a *learning phase*. During an *estimation and prediction phase*, the joint trajectories are measured, remaining parameters are estimated and the future evolution of the joint trajectories is predicted.

The following requirements are imposed on the gait prediction module: it should model the *variations* in human gait and adapt itself to the *velocity and timing* along the joint trajectories (e.g. walking faster/slower, longer/shorter stride length, longer/shorter stance phase, etc.); it should predict in *real-time* the future evolution of the gait trajectories (i.e. the model is *generative*); it should give an *estimate of the uncertainties* on this prediction; and it should measure the *consistency* of the current gait trajectories with the learned model.

A lot of research groups have already tackled part of this problem. For example, Aoyagi [3] predicts the timing in real-time, but starts from a single reference trajectory and does not provide uncertainty or consistency information. The method of Troje [4] [5] focuses on analysis and synthesis of movements and their variations, but does not allow phase/timing or on-line estimation, and is directly based on markers. Other work focuses completely on the analysis in a clinical context, e.g. Federolf [6]. Cheng [7] uses hidden Markov models and Gaussian Processes Latent variables models to analyze human gait. Since these models are mainly for the purpose of human identification, they are not generative.

The contribution of this paper is to provide a statistical modeling method that satisfies all of the above requirements. Analysis tools are provided to optimize the learned model towards its predictive capabilities. Parameters for the estimation/prediction phase are either determined out of the example data set during the learning phase or have a clear physical meaning.

II. A GAIT PREDICTION MODULE

This section describes a gait prediction module and how it fits in a larger control scheme for a lower-limb exoskeleton. Fig. 1 shows an overview of the system. Using a data set of nominal gait patterns, an initial learning phase builds up a model of gait trajectories. The gait is described in terms of the hip, knee and ankle joint angles in the sagittal plane. While measuring the current joint angles, this model is able to predict, in *real-time*, the further evolution of the *joint angle trajectories*, together with a prediction of the *uncertainty* of these estimates as well as the *velocity and timing* along the trajectory. The gait prediction module is able to compute a *consistency* between the measured joint angles and the learned model. In a constraint-based control framework [8], weighted and prioritized constraints describe the control tasks, and at each time step, an optimization problem is solved to optimally satisfy the given control tasks. A prediction of the evolution of the joint angles, together with the estimated uncertainty of this prediction, is used to formulate impedance constraints on the motion of the exoskeleton. The velocity and timing along the trajectory is used to enable additional constraints at certain phases of the gait. These constraints further assist the user, e.g. by enforcing ground clearance. Multiple prediction modules can be simultaneously active. Based on the consistency measure, the supervisor selects the most appropriate prediction module and adapts the priorities and weights of the corresponding constraints. This paper elaborates on the lightly colored blocks in fig. 1.

All authors are with the Department of Mechanical Engineering, KU Leuven, Belgium. All authors gratefully acknowledge the financial support by a strategic basic research grant from the Flemish agency for Innovation by Science and Technology for the MIRAD project (IWT-SBO 120057) and applied biomedical research grant IPSA (IWT-TBM 060799). Corresponding author: erwin.aertbelien@kuleuven.be

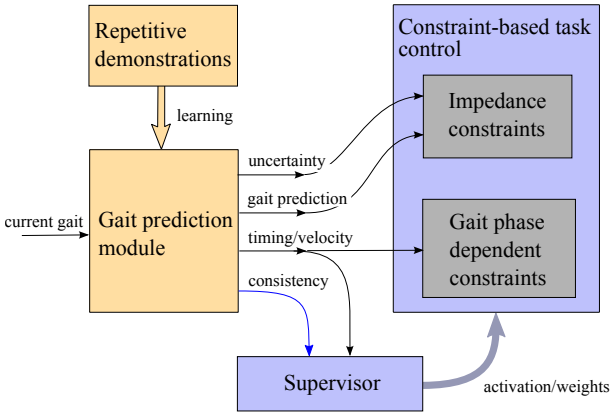

Fig. 1. The use of a gait prediction module for the control of a lower-limb exoskeleton. The lightly colored blocks are the main topic of this paper.

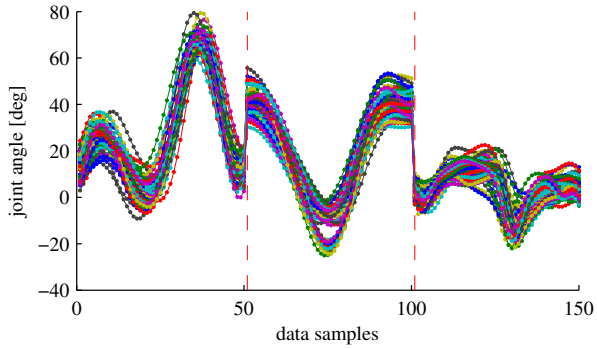

Fig. 2. The trajectory in joint space is correlated over time. Data for 54 gait cycles is shown. Samples 1 to 50 correspond to the knee angle, samples 51 to 100 correspond to the hip angle, samples 101 to 150 correspond to ankle angle.

III. METHODOLOGY

A gait prediction module measures the hip, knee and ankle joint unilaterally in the sagittal plane. Typical human gait is modeled and predicted. However, the same method could be used for more or less joints, using unilateral or bilateral measurements, and for different types of motion. It is assumed that, during the learning phase, the joint trajectories can be segmented using the start and end of the trajectories or a using periodic event. The case of human gait over a flat and horizontal surface is further elaborated, the right side of the exoskeleton is measured, and the segments are determined by foot strike of the right leg.

A. Segmentation, time normalization and resampling

A joint trajectory $\boldsymbol{f}(t)$ is decomposed into a *time normalized trajectory* $\boldsymbol{f}^*(s)$:

$$\boldsymbol{f}(t) = \boldsymbol{f}^*(s(t)) \tag{1}$$

and a *time normalization function* $s(t)$:

$$s = (t - t_0)/T_s, \tag{2}$$

where t_0 corresponds to the time offset, s to the path variable, and T_s to the stride time. $\boldsymbol{f}^*$ corresponds to the commonly used time-normalized curves in gait analysis. $\boldsymbol{f}^*$ is composed of f_k^*, f_h^*, and f_a^*, respectively the knee, hip and ankle components.

For the learning phase, the time normalization parameters (T_s and t_0) and the time normalized trajectory $\boldsymbol{f}^*$ are considered separately. A learning data set of M time normalized and sampled paths is represented by the $M \times d$ matrix:

$$\bar{\boldsymbol{F}} = \begin{bmatrix} \vdots & & \\ f^*_{k,j}(s_1) \cdots f^*_{k,j}(s_n) & f^*_{h,j}(s_1) \cdots f^*_{h,j}(s_n) & \cdots \\ \vdots & & \end{bmatrix} \tag{3}$$

with trial number $j = 1 \ldots M$. $\bar{\boldsymbol{f}}_j$ corresponds to the jth row of $\bar{\boldsymbol{F}}$. The joint angles are resampled such that there are $n = 50$ samples per joint angle ($d = 150$). Fig. 2 shows a plot of a recorded data set. Each curve corresponds to a row in $\bar{\boldsymbol{F}}$.

B. Modeling the normalized joint trajectory

The statistics of $\bar{\boldsymbol{f}}$ are modeled using a statistical model with continuous latent variables. This model is learned using the data set $\bar{\boldsymbol{F}}$ and results in a concise description of $\bar{\boldsymbol{f}}$ in function of the vector $\mathbf{x}$ containing m latent variables, where m is a small number (e.g. 3..7). During execution of the exoskeleton controller, these latent variables are continuously estimated using measurements of the current joint angles. The learned model is then used to predict the further evolution of the joint paths.

Fig. 2 shows that the variations in the normalized trajectories are highly correlated for different values of s and different joints. Chao [9] also indicated this. The modeling method should be able to take into account these correlations over a larger time horizon. Existing statistical representation of trajectories such as Gaussian Mixture Models (GMM) [10] or Stable Estimator of Dynamical Systems (SEDS) [11] only take into account more localized correlations.

Therefore, this paper takes a different approach where the variations on top of the nominal trajectory are modeled using a linear latent variable model. Fig. 3 depicts the following latent variable model:

$$\bar{\boldsymbol{f}}^T = \bar{\boldsymbol{H}}^* \boldsymbol{x}^* + \bar{\boldsymbol{b}}^* + \bar{\boldsymbol{\epsilon}}. \tag{4}$$

$\bar{\boldsymbol{H}}$ is a $150 \times m$ matrix containing m *modes*, $\boldsymbol{x}$ is a $m \times 1$ vector of latent variables, $\bar{\boldsymbol{b}}^*$ is a 150×1 vector containing the averaged time-normalized gait trajectories. $\bar{\boldsymbol{\epsilon}}$ is a 150×1 Gaussian distributed vector:

$$\bar{\boldsymbol{\epsilon}} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 \boldsymbol{I}\right). \tag{5}$$

The latent variables $\mathbf{x}^*$ are distributed as $\mathcal{N}(\mathbf{0}, \mathbf{I})$. This model results in the following distribution for $\bar{\boldsymbol{f}}^T$:

$$\bar{\mathbf{f}}^T \sim \mathcal{N}\left(\mathbf{0}, \bar{\boldsymbol{H}}\bar{\boldsymbol{H}}^T + \sigma^2 \boldsymbol{I}\right). \tag{6}$$

The parameters $\bar{\boldsymbol{H}}$, $\bar{\boldsymbol{b}}^*$ and σ^2 are estimated from the data set using maximum likelihood. This could be done with expectation maximization (EM), but as Tipping [12] proved, this model is closely related to principle component analysis (PCA) and the maximum likelihood estimates can be computed in a similar way. The model in equation (4) is referred

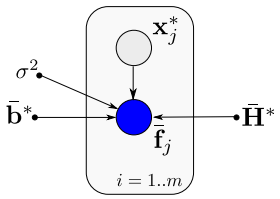

Fig. 3. Probabilistic graphical model (Bayesian network) with continuous variables. See [13] for the graphical notation.

to as *probabilistic principle component analysis* (PPCA). It results in the following expressions for the maximum likelihood estimates $\bar{\boldsymbol{b}}^*_{ML}$, σ^2_{ML} and $\bar{\boldsymbol{H}}^*_{ML}$ of the parameters of model (4):

$$\bar{\boldsymbol{b}}^*_{ML} = \text{mean}(\bar{\boldsymbol{f}}^T) \tag{7}$$

$$\sigma^2_{ML} = \frac{1}{d-m}\sum_{j=m+1}^{d} \lambda_j \tag{8}$$

$$\bar{\boldsymbol{H}}^*_{ML} = \mathbf{U}_m(\Lambda_m - \sigma^2_{ML}\mathbf{I})^{1/2}\mathbf{R}, \tag{9}$$

where $\boldsymbol{U}_m$ are the principal eigenvectors of $\boldsymbol{S}$ with corresponding eigenvalues λ_1,...,λ_m in the diagonal matrix Λ_m, $\Lambda_m = \text{diag}(\lambda_1 \cdots \lambda_m)$, and $\mathbf{R}$ is an arbitrary rotation. $\boldsymbol{S}$ is the sample covariance matrix of $\bar{\boldsymbol{f}}^T$ [12][13]. In contrast to the classical PCA, the columns of $\bar{\boldsymbol{H}}^*_{ML}$ are not normalized, and $\bar{\boldsymbol{H}}^*_{ML}$ can be arbitrarily rotated with the rotation matrix $\boldsymbol{R}$.

Fig. 4 illustrates this decomposition into a number of eigenvectors (also called modes, loadings or coefficients). The gait curves for hip, knee and ankle on the right can be generated from the model when the vector $\boldsymbol{x}$ of latent variables is known.

The number of modes, m, still remains to be determined. Section IV will discuss a cross-validation procedure to determine this number of modes.

Using the model (4), the *function* $\boldsymbol{f}^*(s, \boldsymbol{x}^*) = [f^*_k(s)\ f^*_h(s)\ f^*_a(s)]^T$ can easily be derived:

$$\mathbf{f}^*(s, \mathbf{x}^*) = \begin{bmatrix} h^*_{k,1}(s) & \cdots & h^*_{k,m}(s) \\ h^*_{h,1}(s) & \cdots & h^*_{h,m}(s) \\ h^*_{a,1}(s) & \cdots & h^*_{a,m}(s) \end{bmatrix} \boldsymbol{x}^* + \begin{bmatrix} b^*_k(s) \\ b^*_h(s) \\ b^*_a(s) \end{bmatrix} + \boldsymbol{\epsilon},$$

The functions of the type $h^*_{k,j}(s)$ and $b^*_k(s)$ are one-dimensional interpolation functions for the appropriate elements of the matrix $\bar{\boldsymbol{H}}^*$ and $\bar{\boldsymbol{b}}^*$ of the model (4).

In summary, the latent variable model as a continuous function of s is written as:

$$\boldsymbol{f}^*(s, \boldsymbol{x}^*) = \boldsymbol{H}^*(s)\boldsymbol{x}^* + \mathbf{b}^*(s) + \boldsymbol{\epsilon}. \tag{10}$$

$\boldsymbol{x}^*$ is a m×1 vector $\sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$. $\boldsymbol{f}^*$ and $\boldsymbol{b}^*$ are 3×1 vector functions, $\boldsymbol{H}^*$ is a 3×m matrix functions, and $\boldsymbol{\epsilon}$ is a 3×1 vector of Gaussian white noise.

C. Modeling the time normalization function

The time normalization function (2) is using two parameters, the stride time T_s and the time offset t_0. The statistics for these parameters (mean and covariance) are estimated from the learning data set.

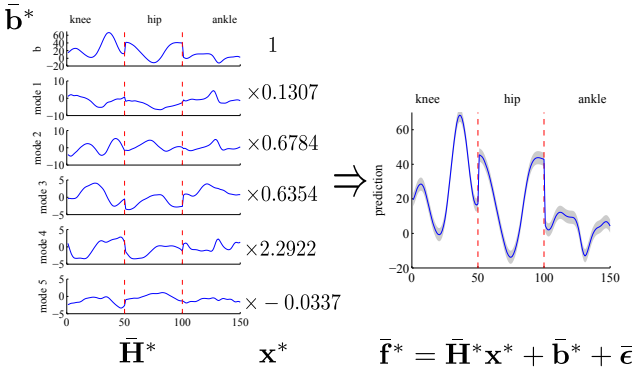

Fig. 4. The latent variable model describes a gait with a limited number of parameters.

D. Real-time estimation of a trajectory

A real-time iterated extended Kalman filter (IEKF) estimates both the value of the latent variables $\boldsymbol{x}^*$ in the model (10) and the time normalization parameters T_s an t_0 using the current measurements of the exoskeleton joints.

The state is described by the latent variables $\boldsymbol{x}^*$, the offset $s_0 = t_0 T_s^{-1}$ and the velocity $v = T_s^{-1}$:

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}^{*T} & v & s_0 \end{bmatrix}^T. \tag{11}$$

The parameters v and s_0 are chosen instead of T_s and t_0 to decrease the non-linearity of the IEKF. The evolution of the state $\boldsymbol{x}$ is expressed using the process model:

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{\rho}_p, \tag{12}$$

with $\boldsymbol{\rho} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q})$, where $\boldsymbol{Q}$ represents the process uncertainty.

For the current measurements of the joint angles, a measurement model is written as:

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}, t) + \boldsymbol{\rho}_m, \tag{13}$$

with $\boldsymbol{\rho}_m \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R})$, where $\boldsymbol{R}$ represents the measurement uncertainty.

The IEKF linearizes this measurement model using a first order Taylor approximation around the current estimate $\boldsymbol{x}_k$:

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}_k, t) + \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x_k}, t} (\boldsymbol{x} - \boldsymbol{x}_k) + \boldsymbol{\rho}_m \tag{14}$$

The measurement function $\boldsymbol{h}$ is determined using the model (10) and the time normalization function in (2):

$$\boldsymbol{h}(\boldsymbol{x}, t) = \boldsymbol{H}^*(vt + s_0)\boldsymbol{x}^* + \boldsymbol{b}^*(vt + s_0) \tag{15}$$

The measurement Jacobian in (14) is computed using:

$$\left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x_k}, t} = \begin{bmatrix} \dfrac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}^*} & \dfrac{\partial \boldsymbol{h}}{\partial v} & \dfrac{\partial \boldsymbol{h}}{\partial s_0} \end{bmatrix} \tag{16}$$

with:

$$\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}^*} = \mathbf{H}^*(vt + s_0) \tag{17}$$

$$\frac{\partial \boldsymbol{h}}{\partial v} = \left(\frac{\partial \boldsymbol{H}^*}{\partial s}\boldsymbol{x}^* + \frac{\partial \boldsymbol{b}}{\partial s}\right) t \tag{18}$$

$$\frac{\partial \boldsymbol{h}}{\partial s_0} = \frac{\partial \boldsymbol{H}^*}{\partial s}\boldsymbol{x}^* + \frac{\partial \boldsymbol{b}}{\partial s} \tag{19}$$

The partial derivative $\frac{\partial \boldsymbol{H}^*}{\partial s}$ is pre-computed during the learning phase using numerical differentiation of $\boldsymbol{H}^*(s)$.

Using the above formulation of the state, the process equation and the measurement equation, an IEKF is formulated to continuously estimate the variations in the gait pattern, the stride time and the time offset.

A few tuning parameters remain. An obvious choice for the measurement uncertainty $\boldsymbol{\rho}_m$ is $\boldsymbol{\epsilon}$ in the model (10). The process uncertainty $\boldsymbol{\rho}_p$ tunes the variability of the gait over a longer time. It models how fast the user can speed up or slow down his walking pattern during gait. The initial estimates for the state $\boldsymbol{x}$ are determined as follows: The model (10) gives the initial value and uncertainty on $\boldsymbol{x}^*$, i.e. $\boldsymbol{x}_0^* \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$. The initial value and uncertainty of v and s_0 are determined from the statistics of the stride time and the time offset.

Once an estimate of $\boldsymbol{x}$ is obtained, the future evolution of the trajectory is given by:

$$\boldsymbol{z}_{\text{pred}} = \boldsymbol{h}(\boldsymbol{x}, t_f) + \boldsymbol{\rho}_{\text{pred}}, \tag{20}$$

where t_f corresponds to a future time and where $\boldsymbol{\rho}_{\text{pred}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\text{pred}})$ corresponds to the uncertainty on the prediction. The prediction uncertainty $\boldsymbol{R}_{\text{pred}}$ is given by:

$$\boldsymbol{R}_{\text{pred}} = \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x},t_f} \boldsymbol{P} \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x},t_f}^{T}, \tag{21}$$

where $\boldsymbol{P}$ corresponds to the uncertainty of the state $\boldsymbol{x}$ as computed by the IEKF.

E. The consistency between model and measurements

As previously explained, it is important to monitor continuously the consistency of the measurements with the model. In this way, the consistency of all running prediction modules can be determined. Based on this, the supervisor module in the controller of fig. 1 will choose the currently active motion and constraints. The consistency of the previously established IEKF is determined using a normalized innovation squared (NIS) test for each of the joint measurements and a summed normalized innovation squared (SNIS) test [14]. The NIS value follows a χ^2-distribution with 1 degree of freedom. The SNIS value for all joint measurements together follows a χ^2-distribution with $3K$ degrees of freedom, where K is the number of NIS values that are summed.

F. Mapping uncertainty to impedance

Uncertainty is mapped to impedance by noting the similarity between the expression for potential energy and for the probability of deviation from the mean value for a Gaussian distribution. The impedance described here mainly consists of stiffness matrix $\boldsymbol{K}$ in joint space for each point along the trajectory. On the one hand, for a deviation $\boldsymbol{\delta}$, and a stiffness matrix $\boldsymbol{K}$, the potential energy α is equal to:

$$\alpha = \frac{1}{2}\boldsymbol{\delta}^T \boldsymbol{K} \boldsymbol{\delta}. \tag{22}$$

On the other hand, if the uncertainty is modeled as a multivariate Gaussian distribution with covariance matrix

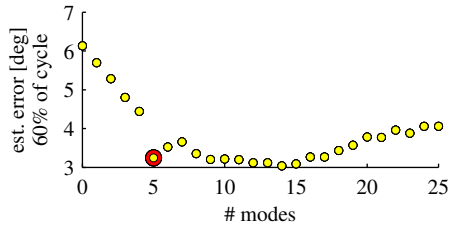

Fig. 5. Choosing the number of modes based on a cross-validation experiment for a training based on sagittal hip, knee and joint angle. The selected number of modes is indicated.

$\boldsymbol{P}_{tr} = \boldsymbol{R}_{\text{pred}}(\boldsymbol{x}, t_f)$, an ellipsoidal region can be defined:

$$\boldsymbol{\delta}^T \boldsymbol{P}_{tr}^{-1} \boldsymbol{\delta} \leq \beta, \tag{23}$$

such that there is a probability p that the deviation is inside the region. β is given by the inverse cumulative χ^2-distribution for p with $k = 3$ degrees of freedom. From (22) and (23), a stiffness matrix $\boldsymbol{K}$ is determined such that a probability p of the measured population results in a potential energy of α or lower:

$$\boldsymbol{K} = \frac{2\alpha}{\beta} \boldsymbol{P}_{tr}^{-1}. \tag{24}$$

This equation gives a mapping between the covariance and a stiffness matrix. This expression allows the exoskeleton controller to give less support (lower stiffness) when the joint trajectory is uncertain.

IV. RESULTS AND DISCUSSION

A. Measurement protocol

A data set for the typical gait of children was established. A standard 3D gait analysis was performed for each included child. Retroreflective markers were then positioned on bony landmarks of the lower limbs as defined by the Plug-In-Gait model of Vicon 1 (Oxford Metrics, Oxford, UK). The children walked on a 10-m walkway at a self-selected speed. Eight infrared Vicon 1 camera's (measuring at 100 Hz) detected the motion of the lower limb markers while the child was walking. Workstation and Polygon software (Oxford Metrics, Oxford, UK) were used to define the gait cycles, to determine the spatio-temporal parameters, and to estimate the joint angles. Each child walked three trials at self-selected speed, three trials at a lower speed and three trials at a higher speed. The gait of seven children was analyzed. Each measurement was visually evaluated for quality. This resulted in measurements of 54 gait cycles for the sagittal motion of the right hip, right knee and right ankle joint. Fig. 2 shows the resulting data.

B. Learning phase

The data set is used to learn the parameters of the model (10) using a maximum likelihood criterion, as explained in section III-B.

A *cross-validation* procedure was used to determine the number of modes m. The data was split up into a training data set and a validation data set containing the gait cycles

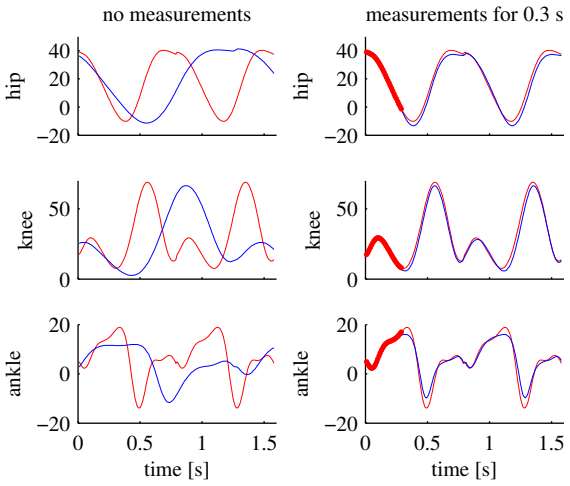

Fig. 6. Estimation of the gait. On the left side, the initial estimate is shown; on the right side, the estimate after $0.3s$. The red curve indicates the actual gait pattern. The thick red curve indicates the measurements. The blue curve shows the actual estimate of the gait trajectory.

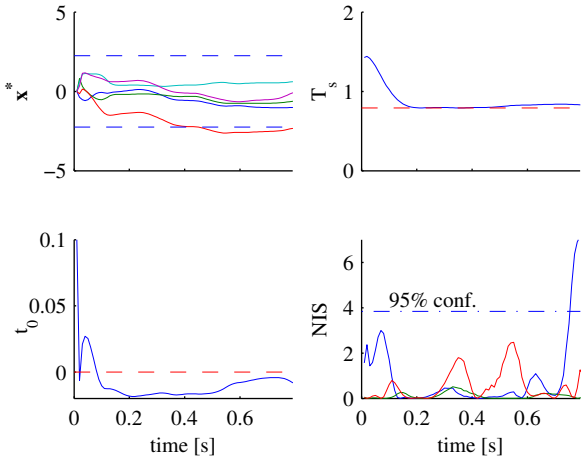

Fig. 7. The evolution of the estimate for the state (latent variables), for the stride time T_s, and time offset t_0. The normalized innovation squared (NIS) for the measurements of the hip (blue), knee (green), and ankle (red).

of one subject. This was repeated for each subject, each time resulting in a different training and validation data set.

For each training/validation pair, the learning method of section III-B is applied for different values of m, ranging from 0 to 25. Using the learned model, the validation data is then used to evaluate the model: The joint measurements from time $0s$ to 60% of the stride time T_s are used to predict the joint trajectories for the whole gait cycle. The root mean squared (RMS) prediction error is determined between the predicted gait cycle and the measured joint trajectory. This resulted in fig. 5. When the number of modes m is very low, the RMS prediction error is high because the model does not capture all variations present in the data set. When the number of modes m is high, the RMS prediction error increases due to overfitting. Based on the fig. 5, $m = 5$ modes are used for the learned model.

C. On-line estimation

One subject is left out of the data set to obtain a training data set that is used to learn a model of the gait patterns. One of the trials of the subject is used to perform the on-line estimation. Fig. 6 shows the results. Initially, at time $t = 0s$, there are no measurements. The initial estimate for the stride time $T_s = 1.43s$ is far away from the actual stride time $T_s = 0.79$. The initial estimate of the time offset $t_0 = 0.1s$, while the actual time offset is $t_0 = 0s$. The left hand side of fig. 6 shows that there is a large difference between the initial prediction of the gait curve and the actual gait curve. The right hand side of fig. 6 shows the prediction results after measuring the three joint angles for $0.3s$ (or approx. 38% of the gait cycle). The estimated gait trajectories converged towards the actual gait pattern. At time 0.3s, the RMS prediction error for the hip, knee and ankle joint is respectively 2.0 degrees, 3.0 degrees and 2.8 degrees.

Fig. 7 shows the time evolution of the estimated parameters. The evolution of the states $\boldsymbol{x}^*$ remains consistent with the distribution $\boldsymbol{x}^* \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$ of the gait pattern model: the 95% two-sided significance interval for each of the elements of $\boldsymbol{x}^*$ is 2.24. The stride time T_s evolves within 200s to the actual stride time $T_s = 0.79$. The time offset evolves very quickly close to the actual value for $t_0 = 0.0$s.

D. Analysis of the performance

To analyze the performance of the gait prediction, a similar cross-validation procedure is used as before. Table I summarizes the results. This table gives the results when no measurements are used, i.e. when just the averaged curves of the training data set are used, when 60% of the gait cycle is measured, and when 100% of the gait cycle is measured. Table I assumes that the stride time T_s and time shift t_0 were known, i.e. it assumes time normalized trajectories. The results of this table were computed using five modes ($m = 5$).

The second column shows the results when the training and validation data set are both equal to the complete data set. The RMS prediction errors are lower since there are no unknown trajectories during the prediction phase. The prediction error after 100% of the gait cycle shows the limits of the linear model used in this paper. Even with 20 modes, this error is still 1.4 degrees.

There was more variability in the ankle joint trajectories. Table I also shows the results when only hip and knee are considered.

As a final remark, this paper considered the gait of children. These children were asked to walk at three different velocities. Furthermore, children have a less developed and systematic gait pattern compared to adults. These factors contributed to the variability of the gait patterns.

In conclusion, the method proposed in this paper can reduce the prediction error by half compared to methods that start from an averaged curve.

TABLE I

RMS ERROR ON PREDICTION

	Cross-validation *[deg]*	Complete data set *[deg]*
Hip, and knee:		
No measurements	5.5	4.8
After 60% of gait cycle	2.5	1.9
After 100% of gait cycle	1.7	1.3
Hip, knee and ankle:		
No measurements	6.1	5.4
After 60% of gait cycle	3.2	2.3
After 100% of gait cycle	2.7	1.9

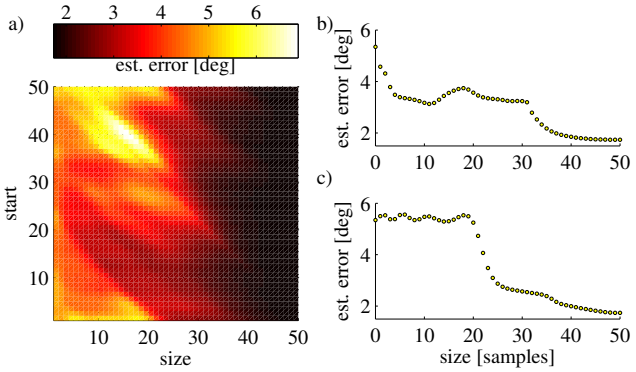

Fig. 8. Prediction of hip and knee joint trajectories: (a) RMS prediction error of the gait estimator for an interval with a given start sample and size, (b) RMS prediction error for intervals starting at sample 20, (c) RMS prediction error for intervals starting at sample 45.

E. Which part of the gait cycle is the most informative?

The prediction method can be used to deduce the interval of the gait cycle that is the most informative for predicting the variations in the gait pattern. This becomes interesting when only hip and knee are measured and predicted. When hip, knee and ankle joints are measured and predicted, information gain is more evenly distributed over the gait cycle.

Using cross-validation, fig. 8a shows a color plot indicating the RMS prediction error for measurement intervals starting at a given sample and with a given size. Figures 8b and 8c show cross-sections of this color plot. Fig. 8b plots the performance of intervals starting at sample 20. Fig. 8c plots the performance of the intervals starting at sample 45.

It can be seen from these plots that the region around sample 20 (40% of the gait cycle) is highly informative. This region falls in the single stance phase of the gait. On the other hand, the region around sample 45-50 (90% - 100% of the gait cycle) is uninformative. This region corresponds to end of the swing phase near foot strike.

V. SUMMARY AND FUTURE WORK

A statistical continuous latent variables model for human gait is developed based on probabilistic principle component analysis. This model takes into account both the nominal joint trajectories of human gait, as well as the variations on these trajectories. It can estimate and predict in real-time the evolution of the joint trajectories, uncertainty of the predictions and the consistency of the model.

A data set of 54 trials of children walking with different velocities is used to validate the proposed method.

Future work will use this gait prediction module on the exoskeleton developed in the MIRAD project with adaptive compliant joints based on the MACCEPA principle[15]. A constraint-based control framework based on [8] will be integrated with the gait prediction module. The mapping explained in section III-F provides a good starting point.

ACKNOWLEDGMENT

The authors want to thank Kaat Desloovere from the Clinical Motion Analysis Laboratory, University Hospital Leuven for the data sets recorded during the IPSA project (IWT-TBM 060799).

REFERENCES

[1] L. Wang, E. H. van Asseldonk, and H. van der Kooij, "Model predictive control-based gait pattern generation for wearable exoskeletons," in *Rehabilitation Robotics (ICORR), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–6.

[2] U. Keller, G. Rauter, and R. Riener, "Assist-as-needed path control for the pascal rehabilitation robot," in *Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–7.

[3] D. Aoyagi, W. Ichinose, S. Harkema, D. Reinkensmeyer, and J. Bobrow, "A robot and control algorithm that can synchronously assist in naturalistic motion during body-weight-supported gait training following neurologic injury," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 3, pp. 387–400, 2007.

[4] N. F. Troje, "Decomposing biological motion: A framework for analysis and synthesis of human gait patterns," *Journal of vision*, vol. 2, no. 5, 2002.

[5] ——, *Retrieving information from human movement patterns*. Oxford University Press, pp. 308–334.

[6] P. Federolf, K. Boyer, and T. Andriacchi, "Application of principal component analysis in clinical gait research: identification of systematic differences between healthy and medial knee-osteoarthritic gait," *Journal of biomechanics*, vol. 46, no. 13, pp. 2173–2178, 2013.

[7] M.-H. Cheng, M.-F. Ho, and C.-L. Huang, "Gait analysis for human identification through manifold learning and hmm," *Pattern recognition*, vol. 41, no. 8, pp. 2541–2553, 2008.

[8] J. De Schutter, T. De Laet, J. Rutgeerts, W. Decré, R. Smits, E. Aertbeliën, K. Claes, and H. Bruyninckx, "Constraint-based task specification and estimation for sensor-based robot systems in the presence of geometric uncertainty," *The International Journal of Robotics Research*, vol. 26, no. 5, pp. 433–455, 2007.

[9] T. Chau, "A review of analytical techniques for gait data. part 1: fuzzy, statistical and fractal methods," *Gait & Posture*, vol. 13, no. 1, pp. 49–66, 2001.

[10] S. Calinon, F. Guenter, and A. Billard, "On learning, representing, and generalizing a task in a humanoid robot," *IEEE Trans. Syst., Man, Cybern. B*, vol. 37, no. 2, pp. 286–298, 2007.

[11] S. Khansari-Zadeh and A. Billard, "Learning stable nonlinear dynamical systems with gaussian mixture models," *IEEE Trans. Robot.*, vol. 27, no. 5, pp. 943–957, Oct 2011.

[12] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 61, no. 3, pp. 611–622, 1999.

[13] C. M. Bishop, *Pattern recognition and machine learning*. Springer New York, 2006.

[14] Y. Bar-Shalom and X.-R. Li, *Estimation and tracking: principles, techniques, and software*. Artech House Boston, 1993, vol. 393.

[15] R. V. Ham, B. Vanderborght, M. V. Damme, B. Verrelst, and D. Lefeber, "Maccepa, the mechanically adjustable compliance and controllable equilibrium position actuator: Design and implementation in a biped robot," *Robotics and Autonomous Systems*, vol. 55, no. 10, pp. 761–768, 2007.